文章编号：1003-1421(2026)07-0194-10
引用格式：张玉召，冀璇，周通，等.基于贝叶斯网络的编组站数据安全风险评估方法研究[J].铁道运输与经济，2026，48(7)：194-203，216.
ZHANG Yuzhao，JI Xuan，ZHOU Tong，et al. Data Security Risk Assessment Methods for Marshalling Stations Based on Bayesian Networks[J]. Railway Transport and Economy，2026，48(7)：194-203，216.

基于贝叶斯网络的编组站数据安全风险评估方法研究

张玉召[1,2]，冀　璇[1]，周　通[3]，徐　嫣[3]，高珅盈泽[1]

(1.兰州交通大学 交通运输学院，甘肃 兰州　730070；2.兰州交通大学 高原铁路运输智慧管控铁路行业重点实验室，甘肃 兰州　730070；3.中国铁道科学研究院集团有限公司 电子计算技术研究所，北京　100081)

摘　要：为应对铁路编组站智能化转型中面临的安全边界模糊、风险量化困难的数据安全治理困境，提出一种基于贝叶斯网络的数据安全风险定量评估方法。通过系统分析编组站业务流程，将数据要素归纳为8大类23子类，识别出数据生命周期中5个环节共15项安全风险要素；构建融合危害程度与风险发生可能性的量化评估模型，并借助贝叶斯网络建模工具构建网络模型，实现节点间依赖关系建模；在数据不完备条件下，采用最大期望估计算法进行参数学习，推理风险的发生概率，最终计算出编组站各类数据在不同风险要素下的安全风险分值。该方法构建了覆盖铁路编组站全业务场景的量化风险评估模型，有效解决了因数据不完备导致的风险量化难题，为差异化的精准防控策略制定提供科学依据，具有重要的理论价值与工程应用前景。

关键词：编组站；数据安全；贝叶斯网络；风险评估；数据生命周期
中图分类号：U29-3　　**文献标识码**：A　　**DOI**：10.16668/j.cnki.issn.1003-1421.20250911006

Data Security Risk Assessment Methods for Marshalling Stations Based on Bayesian Networks

ZHANG Yuzhao[1,2]，JI Xuan[1]，ZHOU Tong[3]，XU Yan[3]，GAO Shenyingze[1]

(1. School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, Gansu, China; 2. Key Laboratory of Railway Industry on Plateau Railway Transportation Intelligent Management and Control, Lanzhou Jiaotong University, Lanzhou 730070, Gansu, China; 3. Institute of Computing Technologies, China Academy of Railway Sciences Corporation Limited, Beijing 100081, China)

Abstract: In response to the data security governance challenges posed by blurred security boundaries and difficult risk quantification during the intelligent transformation of railway marshalling stations, a quantitative data security risk assessment method was proposed based on Bayesian network. Through a systematic analysis of business process, the data elements were categorized into 8 major types and 23 subtypes. 15 security risk factors across 5 stages of the data lifecycle were also identified. A quantitative assessment model integrating severity and risk occurrence probability was constructed. A network model was built using Bayesian network modeling tools to achieve interdependency

收稿日期：2025-09-11
基金项目：中国国家铁路集团有限公司科技研究开发计划课题(P2024X003)

modeling among nodes. With incomplete data, the expectation-maximization algorithm was employed for parameter learning and infer risk occurrence probabilities. Finally, the security risk scores for various data of marshalling stations under different risk factors were calculated. This method establishes a quantitative risk assessment model covering all business scenarios in railway marshalling stations, effectively resolves risk quantification difficulties caused by data incompleteness, provides a scientific basis for formulating differentiated precision prevention and control strategies, and demonstrates significant theoretical value and engineering application prospects.

Keywords: Marshalling Station; Data Security; Bayesian Network; Risk Assessment; Data Lifecycle

0　引言

铁路编组站作为货运网络的核心枢纽，正加速向智能化、数字化转型。基于编组站综合自动化系统、铁路运输管理信息系统、列车调度指挥系统等平台构建的编组站数据中心，其作业模式正经历从传统人工处理向数据驱动的变革，在此过程中必然伴随着大量的数据生成。然而，智能化在提升效率的同时，也放大了数据安全风险，数据在采集、传输、存储、共享环节遭遇泄露或篡改，将对行车安全、个人隐私及企业竞争力构成多层级威胁。此类风险潜藏于数据全生命周期，其安全边界模糊、风险量化困难。因此，如何科学评估编组站数据安全风险，构建精准治理框架，在保障安全的前提下释放数据价值，已成为铁路编组站智能化发展的重要命题。

2021年6月，我国正式颁布《数据安全法》，要求数据所有者需定期开展数据安全风险评估工作。但目前数据安全风险评估工作仍处于探索发展阶段，相关的实施规范和标准尚不成熟。在现有成果中，数据安全能力成熟度模型(Data Security Capability Maturity Mode，DSMM)由国家标准《信息安全技术 数据安全能力成熟度模型》(GB/T 37988—2019)首次提出[1]，该模型为系统评估与提升各组织机构的数据安全能力，提供了一个覆盖数据全生命周期，并融合制度、技术、流程与人员等多维度的管理框架；杨晓琪等[2]在此基础上建立融合管理与技术的数据安全评估模型；王标等[3]构建网络安全等级保护2.0框架，融合传统信息安全风险评估方法，创建一套适用于政府开放数据的安全风险评估模型。

目前针对编组站的研究主要集中在智能化转型的架构与路径上，孙文桥等[4]构建以智能管控为核心、推动编组站管理转型的智能化架构；尹彦博等[5]构建基于多智能体技术的编组站工作组织仿真模型，设计本务机车、调机、车辆和列车智能体及其交互机制，实现编组站作业的动态仿真与优化。在风险评估量化方法方面，传统多采用矩阵法[6]、模糊综合评价法[7]等。这些方法虽在实际工程中得到一定验证，但在处理编组站复杂场景时仍有局限。廖元媛等[8]构建基于贝叶斯网络(Bayesian Network，BN)攻击图的动态风险评估模型，为处理安全事件间的关联性与不确定性提供了参考。在铁路系统数据安全方面，Smith等[9]探讨关键基础设施系统中的信息安全与数据保护问题；陈亚茹等[10]提出数据分级及保护策略，为铁路关键业务领域数据安全治理提供分类与保护的参考框架；陈新柒[11]提出一套针对编组站站区一体化管控平台的数据治理方案，实现了编组站生产作业过程中数据的高度共享；刘珍珍等[12]从数据采集、传输、存储等维度对系统数据安全防护体系进行分析，探究编组站综合自动化系统数据的安全性。

本研究提出了一种基于BN的数据安全风险定量评估方法。该方法首先建立差异化的危害程度量化准则；进而构建覆盖全业务流程的BN模型，通过参数学习优化模型；最后结合危害程度分值与风险发生可能性值，计算安全风险分值，实现对编组站数据安全风险的科学量化与系统评估。

1　编组站全业务数据要素梳理

1.1　数据要素分类

编组站是铁路运输网络中的核心节点，承担列车解体、编组、中转等关键作业，其业务特点为高动态性、强耦合性、安全要求高且复杂度高。基于编组站的业务特点以及铁路数据分级分类标准[13]，编组站数据要素可以分为8类，具体包括：①列车

数据；②货物数据；③作业计划与执行数据；④基础设施数据；⑤视频监控数据；⑥统计与决策数据；⑦用户权限数据；⑧系统配置与日志数据。该分类体系全面涵盖了编组站从生产作业到经营管理各环节的数据资产，为后续开展精准的数据安全风险评估奠定了框架基础。编组站业务数据要素分类如图1所示。

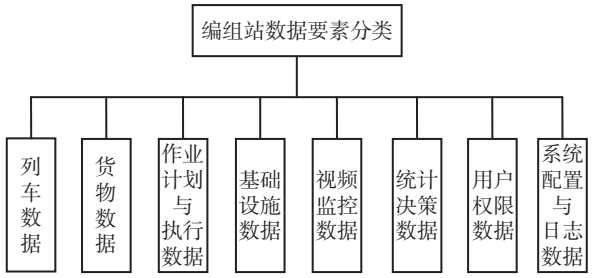

图1　编组站业务数据要素分类

Fig.1　Classification of business data elements for marshalling stations

1.2　数据安全风险要素识别

根据编组站数据生命周期，并参考《数据安全技术 数据安全风险评估方法》(GB/T 45577—2025)[14]中关于数据处理活动的相关规定，从数据采集(R_1)、数据存储(R_2)、数据处理(R_3)、数据传输(R_4)、数据开放共享(R_5)5大环节，梳理出数据安全风险要素15项，编组站数据安全风险要素识别如表1所示。这些风险要素紧密关联编组站作业安全与效率，如

表1　编组站数据安全风险要素识别

Tab.1　Identification of data security risk factors for marshalling stations

序号	数据生命周期	安全风险要素
1	数据采集(R_1)	传感器数据失真(C_{11})
2		超范围采集作业人员隐私数据(C_{12})
3		非法接入采集终端(C_{13})
4	数据存储(R_2)	控制指令存储被篡改(C_{21})
5		数据备份失效(C_{22})
6		日志数据完整性(C_{23})
7	数据处理(R_3)	调度指令篡改(C_{31})
8		自动化决策算法误判(C_{32})
9		越权访问控制指令(C_{33})
10		数据处理延迟导致失效(C_{34})
11	数据传输(R_4)	控制网络数据泄露(C_{41})
12		中间人攻击(C_{42})
13		数据时效性丧失(C_{43})
14	数据开放与共享(R_5)	敏感数据违规共享(C_{51})
15		第三方数据滥用(C_{52})

数据采集环节的传感器数据失真(C_{11})直接危及列车实时追踪与溜放控制的准确性；数据处理环节的调度指令篡改(C_{31})可能引发错误的进路排列与调车作业，进而导致行车事故；而数据开放共享环节的敏感信息违规共享(C_{51})则易造成货主隐私泄露或铁路核心运输数据外泄。

2　编组站数据安全风险评估模型

2.1　编组站数据安全风险评估量化模型

风险评估是识别系统信息安全威胁及脆弱性，并进行定性或定量分析确定风险大小的过程[15]。为实现对编组站数据安全风险的科学量化评估，构建了基于安全风险值R_{ik}的评估模型[14]。模型以数据要素的危害程度和风险发生的可能性为核心指标，综合反映编组站各类数据在不同风险场景下的安全风险水平。R_{ik}的取值范围为[0，1]，数值越大表示风险越高。该模型能够根据实际观测信息和业务场景的变化，动态计算风险发生概率，实现对编组站数据安全风险的定量评估。

$$R_{ik}=\sqrt{\sigma_i\times\rho_{ik}} \tag{1}$$

式中：i为数据要素序号；k为数据风险要素序号；R_{ik}为数据要素i在面临安全风险k时的综合风险分值；σ_i为数据要素i的危害程度分值，反映该数据一旦发生安全风险对行车安全、运输秩序、企业权益与个人权益的影响；ρ_{ik}为数据要素i发生安全风险k的可能性分值，通过BN模型推理获得。

2.2　数据安全风险危害程度

编组站数据安全风险危害程度的量化，依据各数据要素类别自身的敏感程度和重要性来评估。数据一旦发生安全风险，其危害后果主要体现在对行车安全、运输秩序、企业合法权益与个人合法权益的侵害，分为很高、高、较高、中、低5个等级[16-18]。危害程度判断标准主要从4个维度进行评估。①行车安全维度，主要考虑对列车运行控制、信号系统、线路基础设施等系统设施的威胁程度，以及是否可能引发列车冲突、脱轨、冒进等直接危害运输安全的严重后果；②运输秩序维度，主要评估波及范围、社会影响程度、对路网畅通与运输效率的影响、对编组站作为铁路关键枢纽节点功能的影响，以及可能造成的客货运输延误与积压情况；

③企业运营与资产安全维度，主要考虑对编组站及铁路运输企业生产作业连续性、设备设施安全性、商业数据机密性的影响程度，包括运输业务中断时间、应急响应成本及直接经济损失金额；④个人合法权益维度，主要关注客户及作业人员个人信息可能遭受的影响程度，包括是否导致敏感个人信息泄露、是否容易引发诈骗或人身安全风险，以及对个人隐私权的侵害程度。

编组站数据安全风险危害程度的量化，依据《数据安全技术 数据安全风险评估方法》[14]以及饶伟等[13]提出的数据分级思路，将定性的影响程度转化为可量化的风险分值区间，打分区间从高到低分别为[0.8，1]，[0.6，0.8)，[0.4，0.6)，[0.2，0.4)，[0，0.2)。取各阈值的中间值作为数据安全风险危害程度分值[16]，数据安全风险危害程度等级如表2所示。

表2　数据安全风险危害程度等级

Tab.2　Data security risk severity levels

等级	风险危害程度描述	风险危害程度值
很高	一旦发生数据安全风险，对行车安全、运输秩序造成严重危害或特别严重危害，对企业运营与资产安全造成特别严重危害	0.9
高	一旦发生数据安全风险，对行车安全和运输秩序产生危害，对企业运营与资产安全产生严重危害，或对个人信息主体产生特别严重危害	0.7
较高	对行车安全和运输秩序产生有限危害，对企业运营与资产安全产生危害，对个人信息主体合法权益产生严重危害	0.5
中	对行车安全和运输秩序不产生严重危害，对企业运营与资产安全产生一般危害，对个人权益产生危害	0.3
低	对行车安全、运输秩序及企业运营与资产安全几乎不产生危害，对个人权益造成一般危害	0.1

2.3　基于贝叶斯网络的风险发生可能性计算方法

BN是一种基于概率论与图论的不确定性建模工具，能够有效刻画风险要素间的复杂依赖关系，适用于数据不完备场景下的风险概率推理。在编组站数据安全风险评估中，BN为多风险因素的联合建模与推理提供了科学工具，能够有效融合专家经验与历史数据，进行精准的概率推理。研究采用BN对各类数据子类在不同生命周期阶段面临的安全风险发生可能性进行建模与推理。

2.3.1　风险要素影响关系确定

基于1.1节的8类数据要素，进一步可细化出23个数据要素子类。为将这些分类结果嵌入BN模型，将各数据子类定义为数据异常事件(S_1—S_{23})，数据子类转化为数据异常事件对应表如表3所示。

表3　数据子类转化为数据异常事件对应表

Tab.3　Mapping between data subtypes and anomaly events

数据要素类别	数据要素子类	数据异常事件
列车数据	车辆基础信息数据	车辆信息篡改(S_1)
	机车数据	机车数据异常(S_2)
	实时状态数据	实时状态失真(S_3)
	传感器数据	传感器数据失真(S_4)
货物数据	货物属性数据	货物属性违规外泄(S_5)
	关联方信息	关联方信息超范围共享(S_6)
	业务单据数据	业务单据伪造(S_7)
作业计划与执行数据	调车钩计划数据	调车计划失效(S_8)
	实际执行轨迹数据	执行轨迹异常(S_9)
基础设施数据	参数数据	系统参数非法变更(S_{10})
	控制指令数据	控制指令数据失效(S_{11})
	维护记录数据	维护记录伪造(S_{12})
视频监控数据	监控视频数据	监控视频篡改(S_{13})
	生物数据	生物特征数据泄露(S_{14})
统计与决策数据	车站作业指标数据	作业指标失真(S_{15})
	分析报告数据	分析报告传输泄露(S_{16})
	预测数据	预测数据偏差(S_{17})
用户与权限数据	账户信息数据	账户信息篡改(S_{18})
	权限配置数据	权限配置失控(S_{19})
	操作审计数据	操作审计日志缺失(S_{20})
系统配置与日志数据	配置数据	系统配置被非法更改(S_{21})
	运行日志数据	运行日志完整性破坏(S_{22})
	备份元数据	备份元数据损坏(S_{23})

通过深入编组站现场调研，与调度、货运、信息技术等岗位专家进行访谈，梳理编组站数据安全风险要素之间的相互影响关系，明确23个数据子类与15个风险要素(表1中列出)的对应关系，并进一步分析各风险要素之间的因果或依赖关系。通过专家小组讨论，形成风险要素影响关系矩阵，为BN模型结构设计提供科学依据。数据子类与风险要素对应关系如表4所示，编组站数据安全风险要素的影响关系如表5所示。表4仅以列车数据为例展示对应关系，其他类别的数据评估表均采用相同方法构建，表5仅展示数据采集和存储之间的关系，其他环节均采用同样的方法。

表4　数据子类与风险要素对应关系

Tab.4　Correspondence between data subtypes and risk factors

数据类型	数据子类	数据采集			数据存储			数据处理				数据传输			数据开放与共享	
		C_{11}	C_{12}	C_{13}	C_{21}	C_{22}	C_{23}	C_{31}	C_{32}	C_{33}	C_{34}	C_{41}	C_{42}	C_{43}	C_{51}	C_{52}
列车数据	车辆基础信息			√												
	机车相关数据			√												
	实时运行状态	√												√		
	传感器数据	√														

表5　编组站数据安全风险要素的影响关系

Tab.5　Interactions among data security risk factors in marshalling stations

环节	评估指标	数据采集			数据存储		
		C_{11}	C_{12}	C_{13}	C_{21}	C_{22}	C_{23}
数据采集	C_{11}						
	C_{12}			√			
	C_{13}		√				
数据存储	C_{21}						
	C_{22}						
	C_{23}						

2.3.2　贝叶斯网络结构构建

依据上述逻辑，将数据异常事件(S_1—S_{23})与风险要素(C_{11}—C_{52})纳入网络节点，确保风险覆盖的全面性。依据2.3.1节通过专家小组讨论形成的风险要素影响关系矩阵，设置有向边，使用贝叶斯网络建模软件构建出贝叶斯网络风险评估模型如图2所示。该网络结构经由领域专家复核，确保其能真实反映编组站数据安全风险的传递路径。

2.3.3　贝叶斯风险评估模型参数的学习

（1）根节点的先验概率。为确保风险分析逻辑统一、评估过程可量化，模型将所有节点的状态，统一定义为“发生(Yes)”与“未发生(No)”2类。对于根节点，依据历史数据与专家经验设定其先验概率。

（2）中间节点的条件概率。中间节点的条件概率反映了贝叶斯网络中各变量之间的因果依赖关系。在编组站数据安全风险评估的贝叶斯网络模型中，中间节点通常代表数据生命周期中的关键环节或重要影响因素，其条件概率的准确估计对于整体风险推理的科学性和有效性具有重要意义。由于实际编组站数据安全管理中，部分数据难以完整获取，存在缺失或隐变量的情况，因此采用最大期望估计算法(Expectation Maximization，EM)对中间节点的条件概率进行参数学习。该方法通过期望步(E步)和最大化步(M步)的迭代过程，有效解决了因数据不完备导致的参数估计难题，E步基于当前

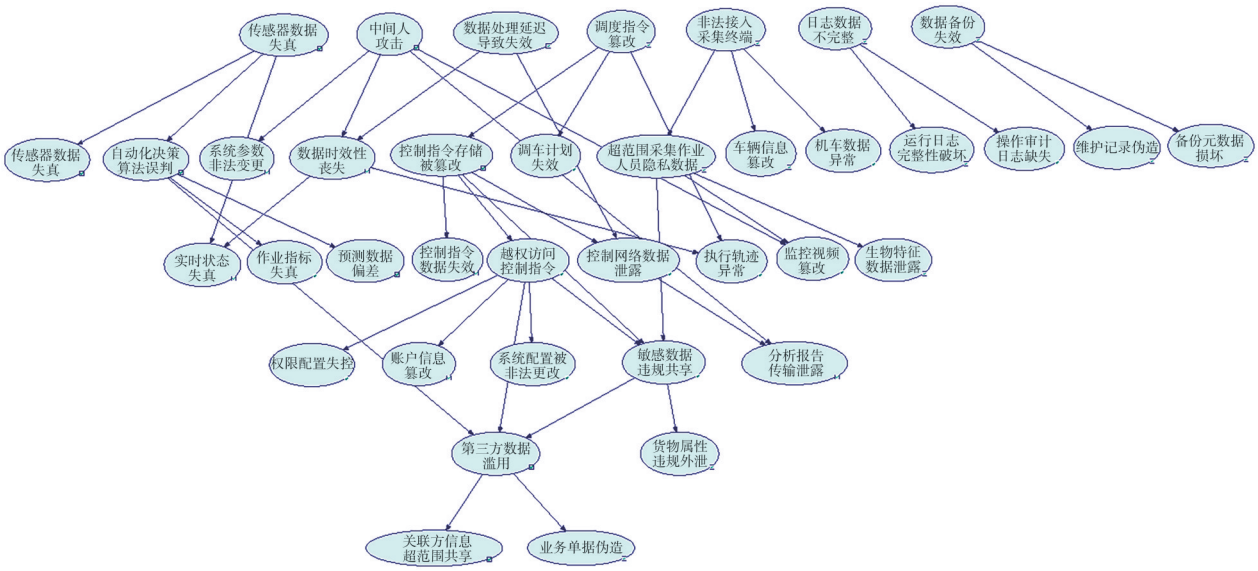

图2　贝叶斯网络风险评估模型

Fig.2　Bayesian network risk assessment model

参数计算隐变量的后验概率分布，M步利用该分布更新网络参数，从而实现对条件概率表的最优估计。EM算法的基本流程包括3个步骤。

①E步：在当前参数下，计算隐变量的后验概率$P(X_i|D_i)$，即在观测数据已知的情况下，推断中间节点可能的状态分布。具体计算公式为

$$P(X_i=1|D_i)=\frac{P(D_i|X_i=1)P(X_i=1)}{P(D_i)} \tag{2}$$

$$P(X_i=0|D_i)=1-P(X_i=1|D_i)=\frac{P(D_i|X_i=0)P(X_i=0)}{P(D_i)} \tag{3}$$

$$P(D_i)=P(D_i|X_i=0)P(X_i=0)+P(D_i|X_i=1)P(X_i=1) \tag{4}$$

式中：X_i表示第i个样本的隐变量，$X_i=\{0,1\}$，表示风险未发生($X_i=0$)或发生($X_i=1$)；D_i表示第i个样本的观测数据；$P(D_i|X_i=1)$和$P(D_i|X_i=0)$表示隐变量在不同状态下观测数据D_i的观测概率；$P(X_i=1)$和$P(X_i=0)$表示隐变量在不同状态下的先验概率；$P(D_i)$表示观测数据的边缘概率，可以通过所有可能的隐变量加权求和来表示。

②M步：利用E步得到的隐变量后验概率，更新BN参数，从而最大化观测数据在隐变量后验分布下的期望似然函数。对于离散节点，参数更新通过计算基于后验概率加权的期望计数并进行归一化来实现。

③迭代：在每次迭代中，E步和M步交替进行，直至模型参数收敛，即模型参数的变化小于预设的阈值时停止迭代。

通过EM算法进行参数学习，得到了一个参数完备的BN模型。该模型的概率推理计算依托于软件实现。

3　编组站数据安全风险定量评估结果与分析

围绕危害程度分值和风险发生可能性分值两大指标，系统展示编组站各类数据要素在不同风险情景下的定量评估结果。通过对评估结果的深入分析，不仅可以揭示高风险数据类型和关键风险环节，还为后续的风险分级、重点防控和安全管理决策提供了坚实的数据支撑和理论依据。

3.1　编组站数据危害程度分值量化结果

根据危害程度分析方法，结合编组站各类数据要素的敏感性、重要性及其对行车安全、运输秩序、企业权益与个人权益的影响，采用德尔菲法的专家打分机制，对8类23种数据要素的安全风险危害程度进行了量化。

为确保评估的全面性与权威性，由涵盖业务、技术与安全管理领域的专家小组，采用德尔菲法通过多轮咨询确定各数据子类的危害程度分值。该方法通过结构化的匿名反馈机制，有效避免了专家会议的从众压力与权威影响，从而系统性地聚合了群体智慧，确保了分值确定过程的科学性与结论的客观性。依据2.2节表2，危害程度区间取值分别为0.9(很高)、0.7(高)、0.5(较高)、0.3(中)、0.1(低)，数据危害程度分值量化结果如表6所示。

表6　数据危害程度分值量化结果

Tab.6　Quantification of data severity scores

数据要素类别	数据要素子类	危害程度等级	危害程度分值
列车数据	车辆基础信息数据	高	0.7
	机车数据	高	0.7
	实时状态数据	很高	0.9
	传感器数据	很高	0.9
货物数据	货物属性数据	较高	0.5
	关联方信息	高	0.7
	业务单据数据	中	0.3
作业计划与执行数据	调车钩计划数据	很高	0.9
	实际执行轨迹数据	很高	0.9
基础设施数据	参数数据	高	0.7
	控制指令数据	很高	0.9
	维护记录数据	中	0.3
视频监控数据	监控视频数据	高	0.7
	生物数据	很高	0.9
统计与决策数据	车站作业指标数据	高	0.7
	分析报告数据	较高	0.5
	预测数据	高	0.7
用户与权限数据	账户信息数据	高	0.7
	权限配置数据	很高	0.9
	操作审计数据	高	0.7
系统配置与日志数据	配置数据	较高	0.5
	运行日志数据	高	0.7
	备份元数据	低	0.1

3.2　编组站数据风险发生可能性分值量化结果

基于BN建模与参数学习，结合实际观测数据和专家知识，利用EM算法对中间节点条件概率进行估计，推理得出各数据要素在不同风险要素下的风险发生可能性分值(ρ_{ik})。在输入部分观测信息后，通过BN，计算各“数据子类–风险要素”组合的发生概率。编组站数据安全风险概率图如图3所示，图3中发生为(Yes)，未发生为(No)。风险发生可能性分值如表7所示。

3.3　编组站数据安全风险值量化结果

根据风险评估模型，对编组站各数据要素面临的安全风险进行综合计算，得到安全风险值R_{ik}以

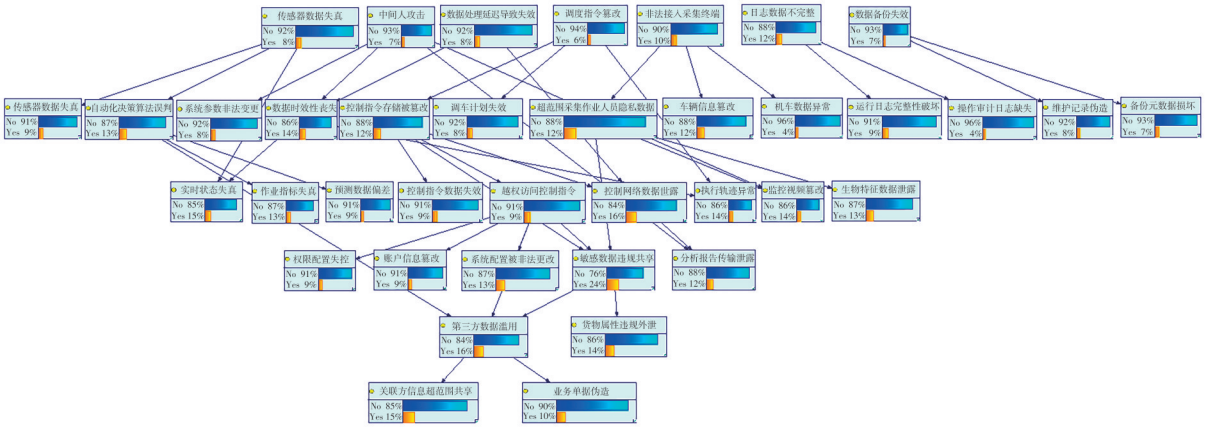

图3　编组站数据安全风险概率图

Fig.3　Overall risk probability for marshalling station data security

表7　风险发生可能性分值

Tab.7　Risk occurrence probability scores

数据异常事件	风险项	风险发生可能性值(ρ_{ik})
车辆信息篡改(S_1)	非法接入采集终端(C_{13})	0.443
机车数据异常(S_2)	非法接入采集终端(C_{13})	0.185
实时状态失真(S_3)	传感器数据失真(C_{11})	0.25
	数据时效性丢失(C_{43})	0.464
传感器数据失真(S_4)	传感器数据失真(C_{11})	0.227
货物属性违规外泄(S_5)	敏感数据违规共享(C_{51})	0.491
关联方信息超范围共享(S_6)	第三方数据滥用(C_{52})	0.5
业务单据伪造(S_7)	第三方数据滥用(C_{52})	0.491
调车计划失效(S_8)	调度指令篡改(C_{31})	0.296
执行轨迹异常(S_9)	调度指令篡改(C_{31})	0.703
	数据时效性丢失(C_{43})	0.166
系统参数非法变更(S_{10})	控制网络数据泄露(C_{41})	0.25
控制指令数据失效(S_{11})	非法接入采集终端(C_{13})	0.166
维护记录伪造(S_{12})	数据备份失效(C_{22})	0.156
监控视频篡改(S_{13})	超范围采集隐私(C_{12})	0.522
	控制网络数据泄露(C_{41})	0.485
生物特征数据泄露(S_{14})	超范围采集隐私(C_{12})	0.166
作业指标失真(S_{15})	日志数据完整性(C_{23})	0.175
分析报告传输泄露(S_{16})	日志数据完整性(C_{23})	0.163
	调度指令篡改(C_{31})	0.25
预测数据偏差(S_{17})	日志数据完整性(C_{23})	0.163
账户信息篡改(S_{18})	越权访问控制指令(C_{33})	0.55
权限配置失控(S_{19})	越权访问控制指令(C_{33})	0.25
操作审计日志缺失(S_{20})	数据备份失效(C_{22})	0.133
系统配置被非法更改(S_{21})	越权访问控制指令(C_{33})	0.466
运行日志完整性破坏(S_{22})	数据备份失效(C_{22})	0.147
备份元数据损坏(S_{23})	数据备份失效(C_{22})	0.14

及风险等级如表8所示。该值综合反映了风险的潜在影响，可用于精准识别高风险环节。

为科学划分风险等级，本方法风险等级的划定严格依据《数据安全技术 数据安全风险评估方法》[14]中表4“数据安全风险评价矩阵”所确立的“风险危害程度–风险发生可能性”二维框架。该标准将风险等级划分为“重大、高、中、低、轻微”5级。为实现从定性评价到定量计算的衔接，将上述定性等级映射为1级至5级风险，并设定了相应的风险值(R_{ik})阈值，基于铁路编组站业务对行车安全具有“低容忍、高后果”的显著特性，对各等级阈值进行了适应性调整，适当提高了中高风险等级的判定门槛。这一设定旨在确保评估结果能够更加敏锐地识别对编组站核心业务构成实质性威胁的风险组合，从而更贴合铁路安全管理的实际决策需求。

风险等级具体划分标准为：1级风险($R_{ik} \geqslant 0.6$)，表明该风险组合一旦发生，后果极为严重且发生可能性较高，必须立即采取坚决的控制措施，优先投入资源进行处置；2级风险($0.5 \leqslant R_{ik} < 0.6$)，需要引起高度关注，应制定明确的风险缓解计划；3级风险($0.4 \leqslant R_{ik} < 0.5$)，需引起中度关注，并纳入重点监控范围；4级风险($0.3 \leqslant R_{ik} < 0.4$)，环节相对可控，通过常规安全管理措施维护，保持持续监测；5级风险($R_{ik} < 0.3$)，风险程度最低，环节可控性强，依托常规安全管理措施维护，同步做好持续监测。

3.4　评估结果分析

基于构建的BN风险评估模型，对编组站8类23项数据子类在15项风险要素下的安全风险值进行了计算与等级划分。一共27个数据安全风险项，其中1级风险4个，2级风险4个，3级风险7个，

表8　安全风险值R_{ik}以及风险等级

Tab.8　Security risk values R_{ik} and risk levels

数据异常事件(S)	风险要素(C)	危害程度(σ_i)	风险发生的可能性(ρ_{ik})	风险值(R_{ik})	风险级别
车辆信息篡改(S_1)	非法接入采集终端(C_{13})	0.7	0.443	0.557	2级风险
机车数据异常(S_2)	非法接入采集终端(C_{13})	0.7	0.185	0.36	4级风险
实时状态失真(S_3)	传感器数据失真(C_{11})	0.9	0.25	0.474	3级风险
	数据时效性丧失(C_{43})	0.9	0.464	0.646	1级风险
传感器数据失真(S_4)	传感器数据失真(C_{11})	0.9	0.227	0.452	3级风险
货物属性违规外泄(S_5)	敏感数据违规共享(C_{51})	0.5	0.491	0.495	3级风险
关联方信息超范围共享(S_6)	第三方数据滥用(C_{52})	0.7	0.5	0.592	2级风险
业务单据伪造(S_7)	第三方数据滥用(C_{52})	0.3	0.491	0.384	4级风险
调车计划失效(S_8)	调度指令篡改(C_{31})	0.9	0.296	0.516	2级风险
执行轨迹异常(S_9)	调度指令篡改(C_{31})	0.9	0.703	0.795	1级风险
	数据时效性丧失(C_{43})	0.9	0.166	0.387	4级风险
系统参数非法变更(S_{10})	控制网络数据泄露(C_{41})	0.7	0.25	0.418	3级风险
控制指令数据失效(S_{11})	非法接入采集终端(C_{13})	0.9	0.166	0.387	4级风险
维护记录伪造(S_{12})	数据备份失效(C_{22})	0.3	0.156	0.216	5级风险
监控视频篡改(S_{13})	超范围采集隐私信息(C_{12})	0.7	0.522	0.604	1级风险
	控制网络数据泄露(C_{41})	0.7	0.485	0.583	2级风险
生物特征数据泄露(S_{14})	超范围采集隐私信息(C_{12})	0.9	0.166	0.387	4级风险
作业指标失真(S_{15})	日志数据完整性(C_{23})	0.7	0.175	0.35	4级风险
分析报告传输泄露(S_{16})	日志数据完整性(C_{23})	0.5	0.163	0.285	5级风险
	调度指令篡改(C_{31})	0.7	0.25	0.418	3级风险
预测数据偏差(S_{17})	日志数据完整性(C_{23})	0.7	0.163	0.338	4级风险
账户信息篡改(S_{18})	越权访问控制指令(C_{33})	0.7	0.55	0.62	1级风险
权限配置失控(S_{19})	越权访问控制指令(C_{33})	0.9	0.25	0.474	3级风险
操作审计日志缺失(S_{20})	数据备份失效(C_{22})	0.7	0.133	0.305	4级风险
系统配置被非法更改(S_{21})	越权访问控制指令(C_{33})	0.5	0.466	0.483	3级风险
运行日志完整性破坏(S_{22})	数据备份失效(C_{22})	0.7	0.147	0.321	4级风险
备份元数据损坏(S_{23})	数据备份失效(C_{22})	0.1	0.14	0.118	5级风险

4级风险9个，5级风险3个，评估结果显示，编组站数据安全风险在分布特征上表现出风险环节相对集中、不同数据类型风险差异显著的特点。

从风险等级分布来看，1级风险4项，其中，执行轨迹异常(S_9)在调度指令篡改(C_{31})风险下的组合风险值最高(0.795)，表明该风险一旦发生将对行车安全与调度效率造成极其严重的直接威胁，必须作为最高优先级的防控对象。该风险项的风险值远高于其他风险项，是由该组合内在的极端危害性与高度敏感的耦合关系所共同决定的。

首先，执行轨迹是列车在站内运行路径的直接反映，一旦异常，将直接引发列车冲突、挤岔、脱轨等灾难性行车事故，其危害程度(σ = 0.9)被定义为“很高”；其次，调度指令是生成执行轨迹的关键源头，两者构成强因果关系，且调度指令本身就是一个脆弱性较高的节点；最后，该风险链路径极短，缺乏有效的缓冲，导致风险一旦发生便无法在中间环节被阻断，即刻转化为安全事故。此案例充分证明，本方法的有效性不仅在于能定量计算风险值，更能通过BN揭示复杂系统内部关键风险的传导与放大机制，这是传统矩阵法难以实现的深度。

其余3项1级风险同样具有高危害程度和高发生可能性的特点，分别为实时状态失真(S_3)在数据时效性丧失(C_{43})风险下的组合；账户信息篡改(S_{18})在越权访问控制指令(C_{33})风险下的组合；监控视频篡改(S_{13})在超范围采集隐私信息(C_{12})风险下的组合。这几类风险一旦发生将直接导致行车实时监控失效、调度决策失误、系统越权访问与恶意操作。这表明，编组站数据安全的核心风险分布于多个关键节点，必须进行系统性防控。

2级风险共计4项，主要涉及车辆信息篡改、关联方信息超范围共享、调车计划失效以及监控视频数据在非法接入采集终端、第三方数据滥用以及数据泄露等风险要素作用下形成中高风险组合，一旦发生将会扰乱运输秩序、侵犯隐私权益，需高度重视并制定针对性控制措施。

从数据生命周期视角分析，数据处理(R_3)与数据传输(R_4)环节风险发生频率较高，尤其是调度指令篡改(C_{31})、数据时效性丧失(C_{43})等风险要素的影响范围广、连锁性强。此外，数据开放与共享(R_5)环节虽风险项较少，但若一旦发生，危害程度显著，尤其涉及个人信息与敏感业务数据时，容易引发法律责任与信誉损失。

从数据类型维度看，列车数据、作业计划与执行数据、控制指令数据等实时性强，与行车安全直接相关，其危害程度普遍较高，一旦结合中高概率的风险事件，极易形成高风险组合。相比之下，统计类、备份类数据虽然发生风险的概率不低，但由于其危害程度较低，整体风险值仍处于中低水平。

综上所述，编组站数据安全治理应聚焦高风险环节和关键数据类型，加强对实时控制类数据在传输与处理环节的保护，严格管控对外数据共享与权限访问，建立基于风险值的动态监测与分级响应机制，实现精准防控与资源优化配置。

4　结束语

研究提出了一种基于BN的铁路编组站数据安全风险评估方法，系统构建了涵盖8大类数据、5个生命周期环节的风险要素体系，建立了融合危害程度与风险发生可能性的量化模型，并借助EM算法实现了不完全数据下的参数学习，最终完成了对编组站数据安全风险的定量评估与等级划分。研究发现编组站数据安全风险呈现明显的多级分布与环节集中特征，高风险事件主要集中于数据处理、数据传输环节，尤其是调度指令篡改、数据时效性丧失等风险项对实时控制类数据及关键支撑类数据构成显著威胁。量化评估结果为编组站实施差异化安全资源配置与精准防控提供了决策依据，即应聚焦高风险环节与关键数据类型，实施精准防控策略。对1级风险须采取实时监控、冗余校验等强控制措施；对2级风险应加强访问控制与合规审查；对中低风险则可纳入常规监测体系，从而实现有限安全资源的最优配置。该方法有效解决了编组站因历史数据缺失而无法量化风险的难题，为铁路编组站数据安全治理提供了技术支持，且构建的BN模型奠定了动态风险评估的理论基础，其网络结构为未来接入实时数据、实现动态风险评估提供了可行的技术框架。未来研究可进一步基于BN模型，重点突破实时数据接入与预警规则生成等关键技术，开发集成化的风险预警平台，最终实现从静态评估

到动态感知、自动预警的跨越，全面提升编组站数据安全治理体系的适应性和可靠性。

参考文献：

[1] 国家市场监督管理总局，国家标准化管理委员会．信息安全技术 数据安全能力成熟度模型：GB/T 37988—2019[S]．北京：中国标准出版社，2019：3-6.

[2] 杨晓琪，白利芳，唐　刚．基于DSMM模型的数据安全评估模型研究与设计[J]．信息网络安全，2021，21(9)：90-95.
YANG Xiaoqi，BAI Lifang，TANG Gang. Research and Design of Data Security Evaluation Model Based on DSMM Model[J]. Netinfo Security，2021，21(9)：90-95.

[3] 王　标，刘兴洋，许　卡，等．政府开放数据的国家安全风险评估模型研究[J]．网络与信息安全学报，2020，6(6)：80-87.
WANG Biao，LIU Xingyang，XU Ka，et al. Research on National Security Risk Assessment Model of Open Government Data[J]. Chinese Journal of Network and Information Security，2020，6(6)：80-87.

[4] 孙文桥，叶　飞，辛姝丽．铁路编组站智能化转型发展内涵及路径探讨[J]．铁道运输与经济，2023，45(8)：75-80.
SUN Wenqiao，YE Fei，XIN Shuli. Connotations and Paths of Intelligent Transformation of Railway Marshalling Stations[J]. Railway Transport and Economy，2023，45(8)：75-80.

[5] 尹彦博，何必胜，陈　鹏，等．基于多智能体的编组站工作组织仿真研究[J]．铁道运输与经济，2025，47(1)：191-199，209.
YIN Yanbo，HE Bisheng，CHEN Peng，et al. Simulation Research on Railway Marshalling Yard Operation Organizations Based on Multi-Agent Technology[J]. Railway Transport and Economy，2025，47(1)：191-199，209.

[6] 国家质量监督检验检疫总局，中国国家标准化管理委员会．信息安全技术 信息安全风险评估实施指南：GB/T 31509—2015[S]．北京：中国标准出版社，2016：22-25.

[7] 张　利，彭建芬，杜宇鸽，等．信息安全风险评估的综合评估方法综述[J]．清华大学学报(自然科学版)，2012，52(10)：1364-1369.
ZHANG Li，PENG Jianfen，DU Yuge，et al. Information Security Risk Assessment Survey[J]. Journal of Tsinghua University (Science and Technology)，2012，52(10)：1364-1369.

[8] 廖元媛，王　剑，田开元，等．基于贝叶斯推理的铁路信号安全数据网信息安全动态风险评估[J]．铁道学报，2020，42(11)：84-93.
LIAO Yuanyuan，WANG Jian，TIAN Kaiyuan，et al. Dynamic Information Security Risk Assessment for Railway Signaling Safety Data Network Based on Bayesian Inference[J]. Journal of the China Railway Society，2020，42(11)：84-93.

[9] SMITH J，RUSSELL S，LOOI M. Security as a Safety Issue in Rail Communications[EB/OL]. (2003-10-15)[2025-12-09]. https：//sc. panda985. com/scholar? hl=zh-cn&q=Security+as+a+safety+issue+in+rail+communications.

[10] 陈亚茹，洪　鑫，张红斌，等．铁路运输调度领域数据分级及保护策略研究[J]．铁道运输与经济，2024，46(2)：134-141.
CHEN Yaru，HONG Xin，ZHANG Hongbin，et al. Data Classification and Protection Strategy in Railway Transportation Dispatching[J]. Railway Transport and Economy，2024，46(2)：134-141

[11] 陈新柒．编组站一体化数据治理方案设计与实现[J]．铁路计算机应用，2023，32(3)：44-49.
CHEN Xinqi. Integrated Data Management Scheme of Marshalling Station[J]. Railway Computer Application，2023，32(3)：44-49.

[12] 刘珍珍，徐永梅．编组站综合自动化系统数据安全技术研究[J]．铁路计算机应用，2017，26(11)：9-11，21.
LIU Zhenzhen，XU Yongmei. Data Security Technology of Marshalling Station Integrated Automation System[J]. Railway Computer Application，2017，26(11)：9-11，21.

[13] 饶　伟，李碧秋，任宸莹，等．铁路数据分类分级保护路径研究[J]．铁道通信信号，2023，59(11)：49-54.
RAO Wei，LI Biqiu，REN Chenying，et al. Research on Classification and Grading Protection Strategy of Railway Data[J]. Railway Signalling & Communication，2023，59(11)：49-54.

[14] 全国网络安全标准化技术委员会．数据安全技术 数据安全风险评估方法：GB/T 45577—2025[S]．北京：中国标准出版社，2025：14-18.

[15] 冯登国，张　阳，张玉清．信息安全风险评估综述[J]．通信学报，2004，25(7)：10-18.
FENG Dengguo，ZHANG Yang，ZHANG Yuqing. Survey of Information Security Risk Assessment[J]. Journal of China Institute of Communications，2004，25(7)：10-18.

[16] 卢　毅，郭新宇，刘　剑，等．自动驾驶道路测试数据安

（下转第216页）